

# EVOLUTION OF LEARNING IN FLUCTUATING ENVIRONMENTS: WHEN SELECTION FAVORS BOTH SOCIAL AND EXPLORATORY INDIVIDUAL LEARNING

Elhanan Borenstein,<sup>1,2</sup> Marcus W. Feldman,<sup>1</sup> and Kenichi Aoki<sup>3,4</sup>

<sup>1</sup>Department of Biological Sciences, Stanford University, Stanford, California 94305

<sup>2</sup>Santa Fe Institute, Santa Fe, New Mexico 87501

<sup>3</sup>Department of Biological Sciences, the University of Tokyo, Hongo 7-3-1, Bunkyo, Tokyo 113-0033, Japan

<sup>4</sup>E-mail: kenaoki@biol.s.u-tokyo.ac.jp

Received August 21, 2007

Accepted December 3, 2007

Cumulative cultural change requires organisms that are capable of both exploratory individual learning and faithful social learning. In our model, an organism's phenotype is initially determined innately (by its genotypic value) or by social learning (copying a phenotype from the parental generation), and then may or may not be modified by individual learning (exploration around the initial phenotype). The environment alternates periodically between two states, each defined as a certain range of phenotypes that can survive. These states may overlap, in which case the same phenotype can survive in both states, or they may not. We find that a joint social and exploratory individual learning strategy—the strategy that supports cumulative culture—is likely to spread when the environmental states do not overlap. In particular, when the environmental states are contiguous and mutation is allowed among the genotypic values, this strategy will spread in either moderately or highly stable environments, depending on the exact nature of the individual learning applied. On the other hand, natural selection often favors a social learning strategy without exploration when the environmental states overlap. We find only partial support for the “consensus” view, which holds that individual learning, social learning, and innate determination of behavior will evolve at short, intermediate, and long environmental periodicities, respectively.

**KEY WORDS:** Cumulative culture, exploratory individual learning, integer phenotype space, overlap of environmental states.

The distinguishing feature of human (*Homo sapiens*) culture as opposed to animal cultures is that it is cumulative (Alvard 2003; Henrich and McElreath 2003; Laland and Hoppitt 2003). Thus, as Tomasello (1999, p. 512) explains, “no single child or group of children could on their own . . . create any version of modern human culture . . . . The most distinctive characteristic of human cultural evolution . . . is the way that modifications to an artifact or a social practice made by one individual or group of individuals often spread within the group, and then stay in place until some future individual or individuals make further modifications—and these then stay in place until still further modifications are made.” Interestingly, cumulateness is perhaps lacking from the lithic

tradition of even our nearest hominid relatives, the Neandertals (Akazawa et al. 1998; Klein 1999). In the case of the Neandertals, innovations are relatively few and far between, so that the Mousterian industry appears to stagnate.

For culture to change cumulatively (i.e., to “evolve” in the human sense), there must be both a source of innovations or novel behaviors and an accurate transfer of information between generations. In fact, the ability to absorb the extant culture and then to build creatively on it must reside in the same organism. The psychological processes that generate innovations and novel behaviors, such as trial-and-error and insight, constitute individual learning. Social learning, which entails the transfer of information

between socially interacting individuals, is the generic term for the psychological processes that support cultural inheritance (Cavalli-Sforza and Feldman 1981; Durham 1991). In social learning, an “observer” adopts (or attempts to replicate) the behavior exhibited by an “exemplar” (Galef 1988; Whiten and Ham 1992; Heyes 1993), e.g., by teaching, imitation, or local enhancement.

Social learning and individual learning are two different strategies, which presumably evolved because of the advantages they conferred in gathering information about the environment. A third source of information, found in the genes, is expressed as innate behavior. Reviews of the factors contributing to the emergence of learning have emphasized the role played by a temporally changing environment (Laland et al. 2000; Richerson and Boyd 2000; Alvard 2003; Henrich and McElreath 2003). These writers agree individual learning, social learning (from the parental generation), and innate determination of behavior are favored by natural selection when environmental changes occur at short, intermediate, and long intervals, respectively (where time is measured in generations). Longer intervals between environmental changes imply longer periods of stasis and hence greater environmental stability (predictability).

Theoretical studies by Boyd and Richerson (1985 chapter 4, 1988), Rogers (1988), and Feldman et al. (1996) provided the motivation, and also some support, for this proposal. More recently, Wakano et al. (2004) and Aoki et al. (2005) showed rigorously that this consensus view is basically true for regularly, and also randomly, changing environments, by simultaneously comparing individual learners, social learners, and “innates” (organisms behaving innately) when they are in direct competition with each other. In particular, they identified a threshold value of the (average) interval between environmental changes, such that the innates are fixed above, but are absent below this threshold. Below this threshold, individual and social learners are at polymorphic equilibrium. Moreover, the equilibrium frequency of social learners increases, and that of individual learners decreases, as the interval becomes longer (see also Rogers 1988; Feldman et al. 1996). In a similar vein, Wakano and Aoki (2006), who investigated a mixed strategy model, showed that the attractive evolutionarily stable strategies (Britton 2003; Doebeli et al. 2004) are a pure innate strategy and a mixed learning strategy. For the mixed learning strategy, dependence on social learning will be stronger, and that on individual learning will be weaker, when the environment is more stable (see also Boyd and Richerson 1985 chapter 4, 1988; Feldman et al. 1996).

In these models, individual and social learning are complementary, so that an increase in the frequency of, or dependence on, one entails a decrease in the frequency of, or dependence on, the other. However, it is more pertinent to assume that the same organism may carry out both individual and social learning, and that these abilities may change independently. In fact, the efflu-

escence of the Upper Paleolithic with its highly innovative and variegated tool set strongly suggests that the individual and social learning abilities of our ancestors both underwent some evolutionary improvement before (or perhaps concurrent with) their departure from Africa. (See Klein 1999 for a similar although not identical view.) This situation can be ameliorated by assuming, as we do in this article, that the propensity for social learning and the strength of individual learning are determined separately by different genes.

Most theoretical investigations of the evolution of social learning in a fluctuating environment have viewed social learning in essentially the same way; a social learner acquires its behavior—or, rather, the information required to express the behavior—from a random member of the parental generation (oblique transmission). This behavior will be adaptive or maladaptive depending on whether the information obtained is appropriate to its current environmental state.

On the other hand, individual learning has been modeled in various ways. The simplest model assumes that an individual learner always achieves the correct behavior (Rogers 1988; Feldman et al. 1996), but suffers an exogenous cost, for example due to mistakes made before the mature behavior is realized. Another model allows for the possibility that an individual learner may adopt either the correct or incorrect behavior (Boyd and Richerson 1988). The extent to which the probability of the former exceeds that of the latter depends on the quality of the environmental information available to the individual learner. In yet a third—the “guided variation” model—the behavioral phenotype is represented by a number (Boyd and Richerson 1985 chapter 4). Each organism acquires its “initial phenotype” by social learning (copying a phenotype from the parental generation), or alternatively, the initial phenotype is determined innately (by its genotypic value). This initial phenotype is then modified by individual learning. The resulting “mature phenotype” is assumed to lie closer on average to the optimal (highest fitness) phenotype in the current environmental state.

The model of individual learning that we adopt in this article is motivated by the guided variation model. However, the most important consequence of individual learning may be that it introduces variation in the mature phenotypes, which as a result are distributed around the initial phenotype. On this view, it is not essential that the center of this distribution be biased toward the optimal phenotype. Hence, we define individual learning to be “exploration” around an initial phenotype, and we describe below how this exploration can be regarded as symmetric (unbiased) or asymmetric (biased).

Another factor missing from some previous theory is that the magnitude of (as opposed to the interval between) environmental changes has been ignored. The models developed by Boyd and Richerson (1985 chapter 4, 1988) are exceptional in this regard, but

in general the subject has not received sufficient attention. We will show in this article that predictions regarding the evolution of individual and social learning are critically affected by whether there is overlap in the states corresponding to the different environments.

An additional body of work that focuses on the trade-off between learning and innately defined behavior is concerned with foraging strategies and takes a purely phenotypic approach (e.g., Rodriguez-Gironés and Vásquez 1997; Eliassen et al. 2007). Although these studies address the same fundamental question as our article, there are several important differences. First, although our interest lies mainly with environmental changes across generations and their effect on the evolution of inherited learning strategies, foraging studies are mostly concerned with the effect of environmental changes (or lack of information) within a single generation and apply optimality analysis to identify the role of learning. Second, these foraging models often have the drawback of not considering the possible interaction (or competition) between different strategies and its potential consequences.

In this article, we define four (generic) strategies, “genetic,” “genetic explorer,” “social learner,” and “social learner explorer,” and develop a new model that permits these strategies to interact and to be in direct competition with each other, rather than being sequestered in separate populations. Our formulation entails that social learning and exploratory individual learning are determined by different genes, making it possible for neither, one of the two, or both to evolve. We investigate how the evolutionary equilibria are affected by environmental periodicity, overlap of environmental states, as well as mutation between the genotypic values.

Before embarking on the technical treatments, it will be useful to provide a brief intuitive explanation of these four strategies. The descriptions differ depending on whether individual learning is assumed to be symmetric or asymmetric. In the former case, the mature phenotype of the genetic (G) strategist is always equal to its genotypic value, and the social learner (S) strategist acquires its mature phenotype without error by pure social learning. Furthermore, the initial phenotypes of genetic explorer (GE) and social learner explorer (SE) strategists are identical to the mature phenotypes of G and S strategists, respectively, but exploratory individual learning yields mature phenotypes that are distributed symmetrically around the initial phenotypes. With asymmetric individual learning on the other hand, the mature phenotype, or distribution of mature phenotypes, of each strategy explained above is biased toward the optimal phenotype in the current environmental state. Hence, all four strategies in this case can on the standard view be regarded as including an individual learning component.

### *The Three-Locus Model*

Assume an infinite population of haploid asexual organisms. This simplifying assumption is shared by all previous theoretical work

(Boyd and Richerson 1985 chapter 4, 1988; Rogers 1988; Feldman et al. 1996; Wakano et al. 2004; Aoki et al. 2005; Wakano and Aoki 2006). The life-history events, to be explained in detail below, are innate determination or social learning, individual learning, natural selection, and reproduction with or without mutation.

The phenotype of an organism is represented by an integer. We follow Boyd and Richerson (1985 chapter 4) in referring to the phenotype after innate determination or social learning as the initial phenotype and the phenotype after individual learning as the mature phenotype. Let  $x$  be the initial, and  $z$  the mature phenotype. The state of the environment,  $y$ , is measured on the same scale as the phenotype. By environmental state  $y$ , we mean that an organism whose mature phenotype lies between  $y - w$  and  $y + w$  ( $w \geq 0$ ) survives to reproduce (has fitness 1). Those organisms whose mature phenotypes lie outside this range leave no descendants (have fitness 0), either genetic or cultural. The fitness function is rectangular, where  $w$  may appropriately be called the environmental tolerance. Hence, all phenotypes within this closed interval  $[y - w, y + w]$  have equal fitness. We will refer to  $y$  as the environmental optimum, although the fitness function is in fact assumed to be flat. Furthermore, we assume that the environment alternates between the two states  $y_1$  and  $y_2$  (where  $y_1 < y_2$ , but  $w$  is the same for both  $y_1$  and  $y_2$ ), with a change occurring every  $l$  generations. Note that the interval between environmental changes is  $l$  generations, but the environmental periodicity or cycle is  $2l$  generations.

Two loci are involved in the expression of the initial phenotype. The first locus determines the genotypic value  $g$ . The second gives the probability  $k$  ( $0 \leq k \leq 1$ ) that the organism will rely on social learning. We assume that social learning is oblique, so that a random surviving member of the parental (previous) generation is copied. When there is no social learning, the initial phenotype  $x$  is simply  $g$ . When social learning occurs, however, the distribution of the initial phenotype  $x$  depends on  $k$ . To formalize these ideas, we define  $\psi_t(x | g, k)$  to be the distribution of the initial phenotype  $x$  among  $g, k$  organisms (organisms with genotypic value  $g$  and probability of social learning  $k$ ) in generation  $t$  and assume that

$$\psi_t(x | g, k) = (1 - k)\delta_{xg} + k\eta_{t-1}(x), \quad (1)$$

where  $\eta_{t-1}(\cdot)$  is the distribution of surviving phenotypes (see below) in generation  $t - 1$ , and  $\delta_{xg}$  is Kronecker's delta.

A third locus determines whether exploratory individual learning occurs. Let  $\lambda(z | x, b)$  be the individual learning function for organisms whose initial phenotype is  $x$  and whose breadth of exploration is  $b$  ( $b \geq 0$ ). For the case of symmetric individual learning we define

$$\lambda(z | x, b) = \frac{1}{2b + 1} \quad \text{on} \quad z = x, x \pm 1, \dots, x \pm b. \quad (2)$$

That is, the mature phenotype  $z$  may take any one of the  $2b + 1$  (in-

teger) values around the initial phenotype  $x$  with equal probability (discrete uniform distribution). When  $b = 0$ , there is no exploration. When  $b > 0$ , equation (2) permits at least two biologically relevant interpretations. It may entail active “exploration” of the “phenotype space.” Alternatively, it may represent passive “error” associated with innate determination or social learning. On the latter interpretation, social learning results in the partial adoption of extant phenotypes with mistakes being made in the copying process.

For asymmetric individual learning we take

$$\lambda(z | x, b) = \frac{1}{2b+1} \text{ on } z = x + 1, (x + 1) \pm 1, \dots, (x + 1) \pm b \quad (3a)$$

if  $x < y$ ,

$$\lambda(z | x, b) = \frac{1}{2b+1} \text{ on } z = x, x \pm 1, \dots, x \pm b \quad (3b)$$

if  $x = y$ , and

$$\lambda(z | x, b) = \frac{1}{2b+1} \text{ on } z = x - 1, (x - 1) \pm 1, \dots, (x - 1) \pm b, \quad (3c)$$

if  $x > y$ . Thus, when the initial phenotype  $x$  does not coincide with the environmental optimum  $y$ , asymmetric individual learning shifts the distribution of the mature phenotypes  $z$  one unit closer to the latter. This implies that some (imperfect) information is available to the organism on where the optimum lies relative to its initial phenotype, which may not always be the case.

Clearly, our model of symmetric individual learning induces behavioral changes that are not biased toward those that work better under current environmental conditions and may therefore deviate from the conventional definition of individual learning. However, we reiterate here our emphasis on the exploratory nature of individual learning as perhaps its most important attribute. In addition, the exploration without bias implied by our model of symmetric learning corresponds to the limiting case of small bias, and as such investigating the consequences of this assumption can serve a useful purpose. We will carefully indicate the sense in which the term individual learning is being used whenever it is not clear from the context.

Thus the model assumes three genetic loci. The first locus defines the innate component  $g$  of the initial phenotype, the second fixes the probability  $k$  of social learning, and the third sets the breadth  $b$  of individual learning. The haplotype of an organism can therefore be represented as  $g, k, b$ . We denote by  $\phi_t(g, k, b)$  the frequency of the  $g, k, b$  haplotype among the newborns of generation  $t$ .

The life cycle is completed when the surviving organisms reproduce asexually. At this stage, mutations may occur at the (first) locus determining the genotypic value (but not at the second or third loci). We assume stepwise mutation at rate  $\mu$ , so that an

organism whose genotypic value is  $g$  will produce three types of offspring with the genotypic values  $g - 1$ ,  $g$ , and  $g + 1$  in the proportions  $\mu/2$ ,  $1 - \mu$ , and  $\mu/2$ , respectively.

In what follows, we consider a simplified version of the above model that limits the number of alleles at the social and individual learning loci to two each. Specifically, we set  $k = 0$  (no social learning) or  $k = 1$  (obligate social learning), and  $b = 0$  (no exploration) or  $b = 1$  (minimal exploration). The assumption  $k = 0$  is equivalent to complete innate determination of the initial phenotype, whereas  $k = 1$  entails that the initial phenotype is acquired by social learning so that the genotypic value is irrelevant. Accordingly, we have four generic strategies corresponding to four generic haplotypes: genetic (abbreviated G) when the haplotype is  $g, 0, 0$ ; genetic explorer (abbreviated GE) when the haplotype is  $g, 0, 1$ ; social learner (abbreviated S) when the haplotype is  $g, 1, 0$ ; and social learner explorer (abbreviated SE) when the haplotype is  $g, 1, 1$ . Note also that when  $b = 1$  the individual learning function (2) or (3) takes each of its three values with equal probability  $1/3$ .

Lastly, at the (first) locus determining the genotypic value, we distinguish two cases. When there is no mutation, we posit two alleles with the genotypic values  $g_1$  and  $g_2$  (where  $g_1 < g_2$ ). However, when there is mutation, we keep track of multiple alleles encoding all genotypic values between  $y_1 - w - 1$  and  $y_2 + w + 1$ .

The recursions are given in Appendix 1.

## Symmetric Individual Learning and No Mutation

Some simple cases can be solved analytically. Assume that individual learning is symmetric (eq. 2) and there is no mutation at the locus determining the genotypic value. Consider two cases.

### CASE 1: OVERLAPPING ENVIRONMENT

Consider overlapping environments so that some (mature) phenotypes are viable in both states  $y_1$  and  $y_2$  (where  $y_1 < y_2$ ). This requires  $y_2 - w \leq y_1 + w$  whence  $w \geq 1$  (because  $w$  can take only integer values). In Appendix 2 we identify an equilibrium at which the post-selection phenotype distribution (phenotype distribution, for short) and the frequencies of the eight haplotypes among newborns (haplotype frequencies, for short) are constant in time, in spite of the fact that the environment is fluctuating. That is,  $\hat{\eta}_t(z^*) = \hat{\eta}(z^*)$  and  $\hat{\phi}_t(g, k, b) = \hat{\phi}(g, k, b)$  for all  $t$ , where the caret indicates equilibrium,  $z^*$  denotes the surviving phenotype,  $g = g_1, g_2, k = 0, 1$ , and  $b = 0, 1$ . It follows that  $z^*$  is distributed between  $y_2 - w$  and  $y_1 + w$  (in the region of overlap between the two states), and  $\hat{\eta}(z^*) = 0$  for  $z^* < y_2 - w$  or  $z^* > y_1 + w$ .

At this equilibrium, it can further be shown that a haplotype either has fitness 1 (no mortality in either environmental state) or fitness less than 1 (deaths occur in one or both states). Clearly, the latter haplotypes are selected out, and a monomorphism or polymorphism of the former will ensue. Specifically,

(1) the G strategy haplotype  $g, 0, 0$  is present at equilibrium if  $y_2 - w \leq g \leq y_1 + w$ ; (2) the GE strategy haplotype  $g, 0, 1$  is present if  $y_2 - w + 1 \leq g \leq y_1 + w - 1$ ; (3) the S strategy haplotype  $g, 1, 0$  is always present; and (4) the SE strategy haplotype  $g, 1, 1$  is always absent. For each of the four strategies,  $g$  is either  $g_1$  or  $g_2$ . It is important to note that the haplotype frequencies are constant at equilibrium, but the values will depend on the initial conditions because all persisting haplotypes have the same fitness. In other words, each such equilibrium is neutrally stable. Extensive numerical work (see below) suggests convergence to this class of equilibria without exception.

The predictions (1)–(4) of the above paragraph are in accord with intuition. Here we briefly explain why the S strategy succeeds whereas the SE strategy fails. Recall that both strategies acquire their initial phenotype by copying a member of the previous generation in proportion to  $\hat{\eta}(z^*)$ . In the case of the S strategy, which depends only social learning, the mature phenotypes will all lie in the region of overlap between the two states,  $y_2 - w$  and  $y_1 + w$ . Hence, the S strategy does not suffer any mortality and has fitness 1 in both states. On the other hand, the SE strategy uses symmetric individual learning (as defined by eq. 2 with  $b = 1$ ) in addition to social learning. Because such exploration causes the phenotype distribution to spread out, the presence of the SE strategy ensures the occurrence of the phenotypes  $z^* = y_2 - w$  and  $z^* = y_1 + w$  at equilibrium. Thus, the initial phenotypes  $y_2 - w$  and  $y_1 + w$ , and consequently the mature phenotypes  $y_2 - w - 1$  and  $y_1 + w + 1$ , are among those produced by the SE strategists in the next generation. The latter are inviable in environmental states  $y_2$  and  $y_1$ , respectively, so that the SE strategy has fitness less than 1 in either state.

**CASE 2: NONOVERLAPPING CONTIGUOUS ENVIRONMENT**

With nonoverlapping but contiguous environments, we have  $y_1 + w + 1 = y_2 - w$  or equivalently  $y_2 - y_1 = 2w + 1$  (the upper bound of state  $y_1$  is smaller than the lower bound of state  $y_2$ , but the distance between the two is 1). In this case, the G and S strategies, which are incapable of exploration ( $b = 0$ ), cannot survive a change of state and are selected out. By contrast, the SE strategy ( $b = 1$ ) has positive viability in both states, and so does the GE strategy ( $b = 1$ ) provided  $g = y_1 + w$  or  $g = y_2 - w$ . This suggests the existence of an equilibrium where the SE strategy, and possibly also the GE strategy, is present.

The following is a description of the features of such an equilibrium; analytical details are given in Appendix 3. First assume the GE strategy is absent, for instance because  $g_1 < y_1 + w < y_2 - w < g_2$ . Then the two haplotypes  $g_1, 1, 1$  and  $g_2, 1, 1$  of the SE strategy clearly have the same fitness, whence their frequencies are constant. In environmental state  $y_1$ , the surviving phenotypes  $z^*$  are distributed between  $y_1 - w$  and  $y_1 + w$ . Importantly, the

phenotype  $z^* = y_1 + w$  is always present (see below) among the surviving adults, so the initial phenotype  $x = y_1 + w$  and consequently the mature phenotype  $z = y_1 + w + 1 = y_2 - w$  necessarily occur in the next generation (specifically, 1/3 of those with initial phenotype  $x = y_1 + w$ ). Hence, when the state changes from  $y_1$  to  $y_2$ , some SE strategists (in fact, those of the mature phenotype  $z = y_1 + w + 1 = y_2 - w$ ) are able to survive the new selection regime. That is, the phenotype distribution in the first generation after the environmental change is concentrated at  $z^* = y_2 - w$ .

Subsequently, if the environment remains in state  $y_2$ , the phenotype distribution will spread out in the interval between  $z^* = y_2 - w$  and  $z^* = y_2 + w$ . Table 1a and 1b gives the equilibrium distributions of the post-selection phenotypes  $\hat{\eta}_t(z^*)$  in generations 1 through 5 for the cases in which the environmental tolerance is  $w = 1$  and  $w \geq 2$ , respectively. Clearly, a mirror-image pattern is repeated when the environment reverts to state  $y_1$ , so that  $\hat{\eta}_t(z^*)$  is cyclic of period  $2l$  in  $t$ .

Let us now investigate the fate of the GE strategy haplotypes  $g_1, 0, 1$  and  $g_2, 0, 1$  that are introduced at low frequencies. Here, we present the argument for the special case of  $w \geq 1, l = 1, g_1 = y_1 + w$ , and  $g_2 = y_2$ . Note that the genotypic value  $g_1$  coincides

**Table 1. Equilibrium distributions of the postselection phenotypes in a contiguous environment with symmetric learning and no mutation. Relative magnitudes rather than probabilities are reported for generations 1 through 5 for the cases in which the environmental tolerance is (a)  $w = 1$  and (b)  $w \geq 2$ . To obtain the phenotype distribution,  $\hat{\eta}_t(z^*)$ , each entry should be divided by the row sum, e.g.,  $\hat{\eta}_3(y_2 - w) = \frac{2}{5}$ . (Symmetric learning entails exploration around the initial phenotype that is not biased toward the optimal phenotype.)**

(a)					
$t$	Phenotype distribution				sum
	$y_2-1$	$y_2$	$y_2+1$		
1	1				1
2	1	1			2
3	2	2	1		5
4	4	5	3		12
5	9	12	8		29

(b)						
$t$	Phenotype distribution					sum
	$y_2-w$	$y_2-w+1$	$y_2-w+2$	...	...	
1	1					1
2	1	1				2
3	2	2	1			5
4	4	5	3	1		13
5	9	12	9	4	1	35

with the upper bound of state  $y_1$ , whereas  $g_2$  lies at the optimum of state  $y_2$  (see above). The general case is analyzed in Appendix 3.

Write  $w_t(g, k, b | y_i)$  for the fitness of haplotype  $g, k, b$  in environmental state  $y_i$  in the  $t$ th generation after a change of state (see Appendix eq. A3 for a formal definition). Because the environment changes every generation ( $l = 1$ ), we need to compute the fitnesses for  $t = 1$  only. The fitness of the SE strategist is,

$$w_1(g, 1, 1 | y_1) = \frac{1}{3} \eta_0(y_2 - w) \tag{4a}$$

$$w_1(g, 1, 1 | y_2) = \frac{1}{3} \eta_0(y_1 + w), \tag{4b}$$

in each of the two states  $y_1$  and  $y_2$  (for both  $g_1 = y_1 + w$ , and  $g_2 = y_2$ ). Here,  $\eta_0(\cdot)$  denotes the postselection phenotypic distribution in the generation before the change of state, and the factor  $1/3$  derives from symmetric individual learning (eq. 2 with  $b = 1$ ). Because the GE strategy is rare, Table 1a and 1b assures us that  $0 < \eta_0(y_2 - w) \leq 1$  and  $0 < \eta_0(y_1 + w) \leq 1$ . Hence, the product of the fitnesses over one environmental cycle ( $2l = 2$  generations) is positive and bounded above by

$$\tilde{w}_{SE}(g, 1, 1) \leq \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}. \tag{5}$$

On the other hand, the fitnesses of the two GE strategy haplotypes in the two states are

$$w_1(g_1, 0, 1 | y_1) = \frac{2}{3}, \tag{6a}$$

$$w_1(g_1, 0, 1 | y_2) = \frac{1}{3}, \tag{6b}$$

$$w_1(g_2, 0, 1 | y_1) = 0, \tag{6c}$$

$$w_1(g_2, 0, 1 | y_2) = 1. \tag{6d}$$

Hence, the products of the fitnesses over one environmental period are

$$\tilde{w}_{GE}(g_1, 0, 1) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9} \tag{7a}$$

and

$$\tilde{w}_{GE}(g_2, 0, 1) = 0 \cdot 1 = 0 \tag{7b}$$

for the two haplotypes.

Thus, in this special case, the GE strategy haplotype  $g_1, 0, 1$  will invade whereas haplotype  $g_2, 0, 1$  will immediately disappear. Moreover, a monomorphic equilibrium in the haplotype  $g_1, 0, 1$  will eventually be reached, replacing the polymorphism in the SE strategy haplotypes.

More generally, as we show in Appendix 3, fixation of the SE strategists is unstable to invasion by the GE strategists of

the appropriate haplotype if the interval between environmental changes  $l \leq 3$ , and it is stable if  $l \geq 5$ . When  $l = 4$ , analysis is uninformative, because the product of the fitnesses over one environmental cycle of length  $2l$  is the same for both strategies. However, numerical work (see below, in particular Table 2c) suggests stability in this case. Numerical work also suggests that the GE strategists will be fixed when  $l = 2$ .

To reiterate, the GE strategy with an innate component can invade when the environment is changeable ( $l \leq 3$ , and moreover go to fixation when  $l \leq 2$ ), whereas the SE strategy that depends partly on social learning is uninvadable in a more stable environment ( $l \geq 4$ ). Although this pattern is at first glance an exact reversal of what might be expected on the consensus view, we will argue later (after reporting numerical results concerning the effect of mutation) that it is in fact consistent. Here, we briefly

**Table 2. Equilibrium properties for four distinct environment ( $y$ ) configurations. *Opt:*  $g$  values are identical to the environmental optima (i.e.,  $g_1=y_1$  and  $g_2=y_2$ ). *Bor:*  $g$  values coincide with the environmental borders (i.e.,  $g_1=y_1+w$  and  $g_2=y_2-w$ ). See the header of each table for specific  $g$  values. *Cycle:* The haplotype/phenotype distribution is cyclic. *Const:* The haplotype/phenotype distribution is constant. The gray cells represent the surviving strategies in each scenario. In cases in which a surviving strategy also includes exploration (i.e., GE or SE), the cell is also marked with hatching for clarity.**

(a)

Wide Overlap ( $y_1=500, y_2=501, w=1$ )  
*Opt:*  $g_1=500, g_2=501$  (the *Bor* case is identical)

Learning	$\mu$	$g$	$l$	G	GE	S	SE	Haplotype Dist	Phenotype Dist
Symmetric	0	<i>Opt</i>	1-1000					<i>Const</i>	<i>Const</i>
	0.02	<i>Opt</i>	1-1000					<i>Const</i>	<i>Const</i>
Asymmetric	0	<i>Opt</i>	1-1000					<i>Const</i>	<i>Cycle</i>
	0.02	<i>Opt</i>	1-1000					<i>Const</i>	<i>Cycle</i>

(b)

Narrow Overlap ( $y_1=500, y_2=502, w=1$ )  
*Opt:*  $g_1=500, g_2=502$ ; *Bor:*  $g_1= g_2=501$ ;

Learning	$\mu$	$g$	$l$	G	GE	S	SE	Haplotype Dist	Phenotype Dist
Symmetric	0	<i>Opt</i>	1-1000					<i>Const</i>	<i>Const</i>
		<i>Bor</i>	1-1000					<i>Const</i>	<i>Const</i>
	0.02	<i>Opt</i>	1-1000					<i>Const</i>	<i>Const</i>
Asymmetric	0	<i>Opt</i>	1-1000					<i>Const</i>	<i>Cycle</i>
		<i>Bor</i>	1-1000					<i>Const</i>	<i>Cycle</i>
	0.02	<i>Opt</i>	1-1000					<i>Const</i>	<i>Cycle</i>

Continued

**Table 2. Continued.**

(c)

Contiguous ( $y_1=500, y_2=503, w=1$ )  
*Opt:*  $g_1=500, g_2=503$ ; *Bor:*  $g_1=501, g_2=502$ ;

Learning	$\mu$	$g$	$l$	G	GE	S	SE	Haplotype Dist	Phenotype Dist
Symmetric	0	<i>Opt</i>	1-1000					Const	Cycle
			1-2					Cycle	Cycle
		<i>Bor</i>	3					Cycle	Cycle
			4-1000					Const	Cycle
	0.02	<i>Opt</i>	1-2					Cycle	Cycle
			3					Cycle	Cycle
		<i>Bor</i>	4-11					Const	Cycle
			12					Cycle	Cycle
0.02	<i>Opt</i>	13-1000					Cycle	Cycle	
		1					Const	Cycle	
	<i>Bor</i>	2-1000					Const	Cycle	
		1-1000					Const	Cycle	
0.02	<i>Opt</i>	1-106					Cycle	Cycle	
		107-144					Cycle	Cycle	
	<i>Bor</i>	145-1000					Cycle	Cycle	
							Cycle	Cycle	

(d)

Distant ( $y_1=500, y_2=504, w=1$ )  
*Opt:*  $g_1=500, g_2=504$ ; *Bor:*  $g_1=501, g_2=503$ ;

Learning	$\mu$	$g$	$l$	G	GE	S	SE	Haplotype Dist	Phenotype Dist
Symmetric	0	<i>Opt</i>	1-1000					Extinct	
			<i>Bor</i>	1-1000				Extinct	
	0.02	<i>Opt</i>	1					Extinct	
			2-1000					Cycle	Cycle
Asymmetric	0	<i>Opt</i>	1-1000					Const	Cycle
			<i>Bor</i>	1-3				Cycle	Cycle
			4-1000					Const	Cycle
	0.02	<i>Opt</i>	1					Cycle	Cycle
			2-199					Cycle	Cycle
			200-1000					Const	Cycle

note that the key to reconciliation is to equate the G strategy with baseline innate determination. Then, GE represents an “evolved” strategy that has acquired the capacity for symmetric individual learning (i.e., exploration), whereas SE represents a mixed symmetric individual and social learning strategy in which the innate component has been suppressed.

### Asymmetric Individual Learning and No Mutation in a Contiguous Environment

The assumption of asymmetric individual learning renders analysis more difficult. Suppose the initial phenotype  $x$  has a lower value than the environmental optimum  $y$  ( $x < y$ ). Then, equation

(3a) entails that, for a nonexplorer ( $b = 0$ ), the mature phenotype will be  $z = x + 1$ . On the other hand, the mature phenotype of an explorer ( $b = 1$ ) whose initial phenotype is  $x$  will be distributed over  $z = x, x + 1$ , and  $x + 2$  with equal probability of 1/3.

The equilibrium at which the SE strategy is fixed is of special interest, because at this equilibrium both social and (biased as well as exploratory) individual learning abilities are present in the same organism, allowing culture to change cumulatively. Here we briefly report results regarding the stability of this equilibrium to invasion by the S strategy. As shown in Appendix 4, when the environmental tolerance is  $w = 1$  and the interval between changes of state is  $l = 1$ , the product of the fitnesses over one cycle of duration  $2l = 2$  is greater for the rare S strategists than for the dominant SE strategists. Hence in this case the S strategy invades, and goes all the way to fixation as shown in Table 2c. On the other hand, if we assume  $w = 1$  and  $l \geq 2$ , the product of the fitnesses is the same for both strategies. Nevertheless, numerical work (Table 2c) suggests that the S strategists will be selected out, although slowly.

In addition, when the environmental tolerance is  $w = 2$ , the SE strategy is stable to invasion by the S strategy provided the interval  $l$  is sufficiently long. Appendix 4 gives a sketch of the proof.

### Numerical Work

As demonstrated in the previous sections, there are several factors that affect the dynamics and equilibria in our model. We have shown that the environmental periodicity  $2l$  and the magnitude of environmental changes (e.g., overlapping versus contiguous environmental states) markedly affect the evolution of individual and social learning. In this section we explore numerically the equilibria of the haplotype frequencies and the phenotype distribution in various environments and for a range of parameter values. We validate the analytical findings obtained above and examine the equilibria under more complicated scenarios (e.g., when the genotypic value  $g$  can mutate) where a formal mathematical analysis has not been feasible.

The results presented below have all been obtained according to the following protocol. Initial frequencies of all possible haplotypes were set equal. Initial phenotype frequencies were uniformly distributed across all viable phenotypes in the pre-change environment. The recursion equations (see Appendix 1) were then iterated until equilibrium was reached. Specifically, simulations were terminated when the difference in haplotype (and phenotype) frequencies between two adjacent cycles (each of length  $2l$ ), summed over all haplotypes (and phenotypes) and over all generations in the cycle, fell below  $10^{-7}$ .

Once equilibrium was reached, we examined the surviving haplotype and phenotype frequencies. Haplotypes with mean frequency lower than  $10^{-4}$  were considered absent. Additional

simulations have confirmed that these low-frequency haplotypes are indeed subject to weak selection and do eventually disappear. We also examined the haplotype and phenotype distributions to determine whether the equilibrium frequencies are constant or experience a cycle.

The results obtained for a large set of simulation runs are summarized in Table 2a, 2b, 2c and 2d; we have considered four distinct environments, termed *Wide-Overlap*, *Narrow-Overlap*, *Contiguous*, and *Distant*, which are characterized by the distance between the two environmental optima,  $y_1$  and  $y_2$  (see also the headers of Table 2a, 2b, 2c and 2d). The environmental tolerance,  $w$ , was set to 1 in all simulations. For each such environment, we examined the effects of the interval between changes of state ( $l$ ), different genotypic values ( $g_1$  and  $g_2$ ), mutation rate ( $\mu$ ), and symmetric versus asymmetric individual learning, as described in the Section “The Three-Locus Model.” In particular, all values of  $l$  between 1 and 25, and selected values of  $l$  up to 1000 were considered. It is evident (see for example Table 2b) that all of these parameters can have a marked effect on the resulting equilibria.

Table 2a and 2b records the equilibria when the environmental states overlap (wide-overlap and narrow-overlap, respectively) and confirm the analytical results obtained in the Section “Symmetric Individual Learning and No Mutation” for the case of symmetric learning and no mutation. Thus, the G strategy haplotype is present at equilibrium if  $g$  is located in the overlapping region; see for example Table 2b, and compare the case where the  $g$  values are equal to the environmental optima ( $g_1 = 500$ ,  $g_2 = 502$ ) with the case where the  $g$  values coincide with the environmental borders ( $g_1 = g_2 = 501$ ). The GE strategy haplotype is not present, as the condition for this to occur (i.e.,  $y_2 - w + 1 \leq g \leq y_1 + w - 1$ ) is not met by our choice of parameters. The S strategy haplotype is always present, and finally the SE strategy haplotype is always absent. Furthermore, as predicted by our analysis, the haplotype frequencies and the phenotype distribution are both constant.

However, it appears that the presence of mutation or the use of asymmetric rather than symmetric learning significantly influences the equilibria. Specifically, when the genotypic values  $g$  can mutate, the G strategists are selected out. This is in accord with intuition as a mutation in a viable  $g$  allele may make it nonviable; consequently when the mutations accumulate, the fitness of G strategists will eventually fall below 1. (In an overlapping environment, there is selection against the mutable G strategists, under both symmetric and asymmetric learning assumptions.) Similarly, asymmetric learning sometimes results in the survival of GE strategists that were absent from the equilibrium in the symmetric learning case.

Table 2c validates our analysis presented in Sections “Symmetric Individual Learning and No Mutation” and “Asymmetric Individual Learning and No Mutation in a Contiguous Environ-

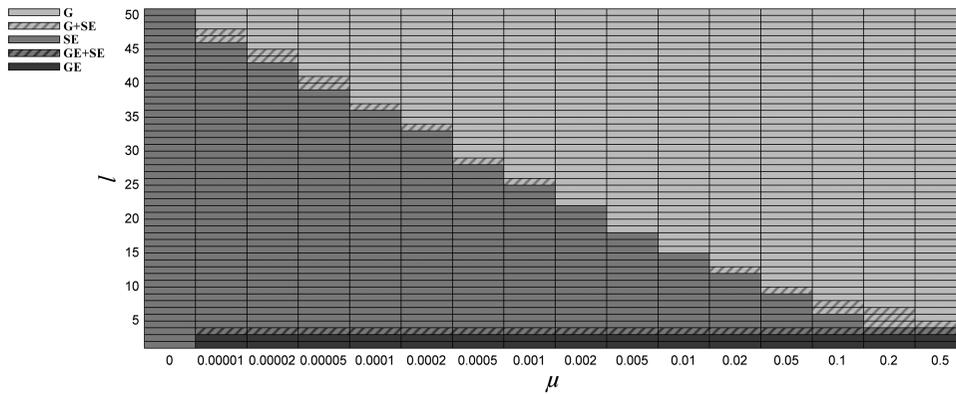
ment” for the contiguous environment without mutation. The effect of  $l$ , the interval between changes of state, is evident in the symmetric learning case. Although the GE strategy spreads in more changeable environments (small  $l$ ), the SE strategy spreads and eventually fixes when the environment becomes more stable ( $l \geq 4$ ). We also see that when organisms apply asymmetric individual learning, an equilibrium at which the SE strategy is fixed may arise for  $l \geq 2$ .

The results recorded in Table 2c for a contiguous environment demonstrate that the presence of mutation,  $\mu > 0$ , and the interval between changes of state,  $l$ , have major effects on the surviving haplotype frequencies. Apparently, with symmetric learning in a contiguous environment, and when mutations in the  $g$  alleles are permitted, three markedly different regimes can be observed as a function of the level of environmental stability. When  $l$  is small (low stability), only GE (genetic explorer) strategists survive. For intermediate  $l$  values, only SE (social learner explorer) strategists survive (see Table 2c for symmetric learning, mutation, and  $4 \leq l \leq 11$ ). And finally when  $l$  is large (relatively stable environment), only G (genetic) strategists survive. Certain  $l$  values that correspond to transition values from one regime to another result in a polymorphism of two strategies. Figure 1 shows the effect of both the mutation rate,  $\mu$  ( $0 \leq \mu \leq 0.5$ ), and the interval between changes of state,  $l$  ( $1 \leq l \leq 50$ ). Note that a transitional polymorphism is observed in most cases (i.e., for most  $\mu$  values) and that the minimum  $l$  value above which the G strategists outperform the SE strategists decreases as the mutation rate increases.

That transitions from GE strategists to SE strategists and then to G strategists occur as  $l$  increases is consistent with the consensus view. Specifically, if we regard the G strategy as the baseline innate determination, then GE represents an “evolved” strategy involving exploratory individual learning, and SE represents a mixed individual and social learning strategy in which the innate component has been suppressed (see also Section 3 “Symmetric Individual Learning and No Mutation”). Thus, these results correspond closely to those of Wakano et al. (2004) and Aoki et al. (2005). We will address the differences from Boyd and Richerson (1985) in the Discussion.

Finally, Table 2d summarizes the numerical results for the distant environment (i.e.,  $y_2 - y_1 > 2w + 1$ ). In some cases when the distance between the two environmental states is too large, there are no haplotypes that can survive an environmental change and the entire population goes extinct.

In addition to the set of surviving haplotypes, equilibria may also differ in the way the haplotype distribution changes over time. In many cases, the resulting equilibria are cyclic (of period  $2l$ ), tracking the environmental cycle. A typical case is portrayed in Figure 2; in the generation immediately following the change of state only one phenotype is present, producing a phenotype



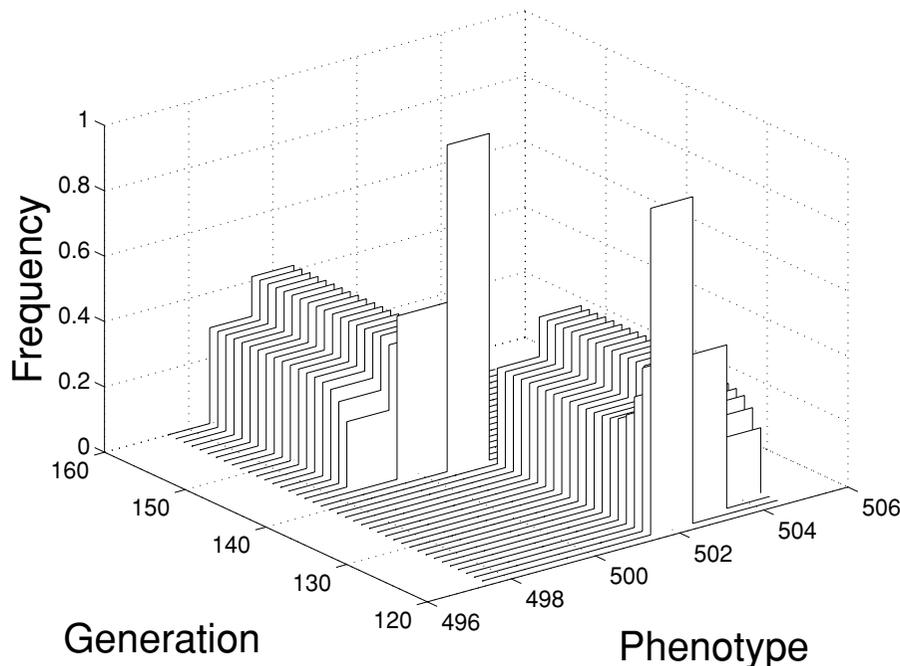
**Figure 1.** Surviving haplotype classes as a function of the interval between changes of state and the mutation rate in a contiguous environment ( $y_1 = 500, y_2 = 503, w = 1$ ), with  $g$  values identical to the environmental optima (i.e.,  $g_1 = y_1$  and  $g_2 = y_2$ ), and symmetric learning. (Symmetric learning entails exploration around the initial phenotype that is not biased toward to optimal phenotype.)

distribution with a single peak. The phenotype distribution becomes broader in subsequent generations until the next environmental change occurs. Figure 3 further illustrates a scenario in which both the phenotype and haplotype distributions are cyclic. (See also Table 1a and 1b.)

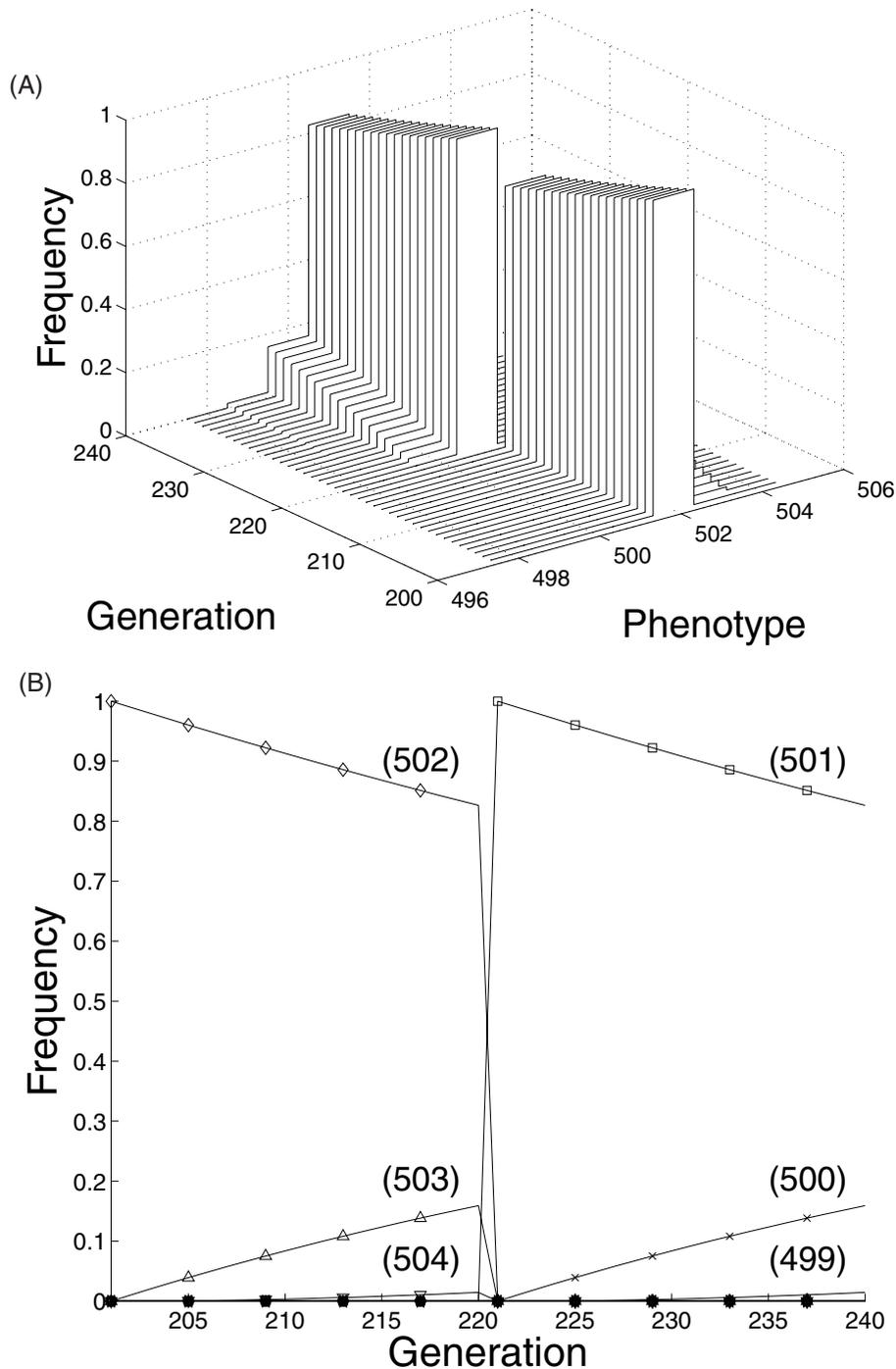
### Discussion

We have described a model in which an organism’s phenotype, represented by an integer, is initially determined innately (by its genotypic value) or by social learning (copying a phenotype from

the parental generation), and then may or may not be modified by individual learning to yield its mature phenotype. Here, individual learning is defined to be exploration, which can be either symmetric or asymmetric around the initial phenotype (unbiased or biased, respectively, toward the optimal phenotype). The environment alternates periodically between two states measured on the same scale as the phenotype, and the surviving mature phenotypes in each state are determined by a rectangular fitness function. The two environmental states may overlap, in which case the same phenotype can survive in both states, or they may be nonoverlapping. This model is novel in that, at least in the symmetric learning



**Figure 2.** The phenotype distribution at equilibrium in a contiguous environment ( $y_1 = 500, y_2 = 503, w = 1$ ), with  $l = 20$ , symmetric learning, no mutation, and innate  $g$  values that coincide with the environment borders (i.e.,  $g_1 = 501$  and  $g_2 = 502$ ). The phenotype distribution over all the phenotypes that exist at the equilibrium is illustrated for each generation within the first cycle after equilibrium was reached (generations 121–160 in this case).



**Figure 3.** The phenotype distribution and the haplotype distribution in a contiguous environment with  $l = 20$ , symmetric learning,  $\mu = 0.02$ , and innate  $g$  values identical to the environmental optima (i.e.,  $g_1 = 500$  and  $g_2 = 503$ ). (A) The phenotype distribution over all the phenotypes that exist at the equilibrium is illustrated for each generation in the first cycle after equilibrium was reached (generations 201–240 in this case). (B) The haplotype distribution of pure innate individuals (i.e., haplotypes of the form  $g,0,0$ ) for different  $g$  values. The distribution is illustrated as a function of the generation for all the generations within the first cycle after equilibrium was reached. The value of the specific  $g$  allele is next to the appropriate curve. All other haplotypes are absent from this equilibrium.

situation, social and individual learning can evolve independently of each other.

We have analytically and numerically investigated the evolutionary fate of four directly competing strategies, which is also

a novel approach. A major result is that a joint social and exploratory individual learning strategy is likely to spread when the environmental states do not overlap. In particular, when the environmental states are contiguous and mutation is allowed among

the genotypic values, this strategy will spread in moderately stable environments when individual learning is symmetric, and in highly stable environments when it is asymmetric. On the other hand, natural selection often favors a social learning strategy without exploration when the environmental states overlap.

The extensive numerical work summarized in Table 2a, 2b, 2c, and 2d and in Figures 1–3 reveals a bewildering variety of possible outcomes, in addition to and including those mentioned in the above paragraph. In the previous sections we have provided analytic or intuitive justification for some of these. Here, we focus on the several new results of greatest theoretical significance for our understanding of the evolution of learning in animals and humans.

First, we address in more detail the question of when selection favors an organism capable of both social and exploratory individual learning, which is a prerequisite for cumulative culture. In other words, we wish to identify the conditions under which the SE strategy will be fixed, or alternatively exist at polymorphic equilibrium. Recall that SE entails the acquisition of the initial phenotype by social learning, which is followed by exploratory individual learning to yield the mature phenotype. Alternatively, SE can be regarded as a social learning strategy that is prone to copying errors.

Table 2a and 2b shows that in the overlapping environment (see illustrations at the top of the tables) SE is always selected out whereas S is often favored (except when a genotypic value coincides with an environmental border). However, when individual learning is asymmetric, we have already noted that the S strategy can be regarded as comprising both individual and social learning, because for an organism adopting this strategy the mature phenotype lies closer to the environmental optimum than the initial phenotype. On this interpretation, selection is favoring an organism capable of both social and biased individual learning (but lacking the ability to explore). There is unfortunately some ambiguity in how individual learning might be defined.

By contrast, Table 2c shows that SE is often fixed in the contiguous environment. In fact, when the genotypic values are not mutable, S is always lost (barring one exceptional case), and SE is fixed except when the environment is highly unstable (or when learning is asymmetric and a genotypic value coincides with the border of an environmental state). When mutation is introduced, the same table shows spread of SE in moderately stable environments when individual learning is symmetric, and in highly stable environments when it is asymmetric. On the alternative interpretation of SE as an error-prone social learning strategy—S entails faithful copying—these results suggest that mistakes made in the copying process may contribute to survival in the contiguous environment.

For completeness, we note from Table 2d that in the distant environment the population goes extinct most of the time if learning is symmetric, and SE can never evolve. With asymmetric

learning we observe a pattern of dependence on the environmental stability similar to the case of the contiguous environment.

In addition, individual learning in our model does not entail the discovery of new behaviors, but only the “rediscovery” of the same behaviors corresponding to phenotypes lying in the range between  $y_1 - w - 1$  and  $y_2 + w + 1$ . This is an artifact of our assumption of two states between which the environment fluctuates. If instead we were to posit an infinite number of states (Feldman et al. 1996; Wakano et al. 2004; Aoki et al. 2005) on the space of integers, with environmental changes being described by, say, a random walk on this space, some behaviors acquired by individual learning would be truly novel. It should also be interesting to adopt a multi-dimensional phenotype space and to investigate the effect of the number of dimensions on the evolution of learning.

Second, we discuss how our findings relate to the prior study by Boyd and Richerson (1985), who compared the relative advantages of genetic and cultural transmission based on their guided variation model. These authors compute the geometric mean fitnesses of two asexually reproducing populations in a stochastically changing environment. In the first population, the organisms use a strategy of innate determination combined with asymmetric individual learning—the “genetic strategy.” In the second, they use a joint social and asymmetric individual learning strategy—the “cultural strategy.” Boyd and Richerson (1985, Figs. 4.11, 4.12, and 4.13) find that low environmental autocorrelations (an unstable environment) favor the genetic strategists—the geometric mean fitness of the first population is higher—whereas intermediate-to-high autocorrelations favor the cultural strategists. Furthermore, at very high autocorrelations, the genetic strategists are again favored. These results are remarkable in that they apparently contradict the consensus view—individual learning, social learning, and innate determination of behavior are favored by natural selection when environmental changes occur at short, intermediate, and long intervals, respectively—and suggest that it may be too simplistic.

However, we have been unable to replicate Boyd and Richerson’s result. Assuming asymmetric individual learning, the equivalent observation within the framework of our model would be that the GE (genetic explorer) strategy, the SE (social learner explorer) strategy, and—once again—the GE strategy are favored by natural selection when environmental changes occur at short, intermediate-to-long, and very long intervals, respectively. Alternatively, the sequence might be G (genetic), S (social learner), and back to G. It can be seen from Table 2a, 2b, 2c, and 2d that this pattern is never realized.

Additional simulations were conducted in which we fixed the breadth of individual learning, so that the number of competing generic strategies was two rather than four—G against S when  $b = 0$ , and GE against SE when  $b = 1$ . Note that these special assumptions more closely match those made by Boyd and Richerson.

In the case of  $b = 1$ , we did observe a transition from GE to SE as the interval between environmental changes increased, but never a transition back to GE, within the range of values considered ( $1 \leq l \leq 1000$ ). Clearly, we cannot rule out the possibility that the GE strategy would be favored again at some higher value of  $l$  (see below).

Nevertheless, we believe that our inability to replicate Boyd and Richerson's result stems from the numerous structural differences between their model and ours. For example, the models differ in whether environmental fluctuations are stochastic or periodic, the fitness function is Gaussian or rectangular, etc.. But these disparities are relatively minor. What really distinguishes our model is the co-occurrence in the same population of various strategies, which compete directly with each other in a frequency-dependent manner. By contrast, Boyd and Richerson assume that the genetic and cultural strategies are evolving independently in two separate populations. Although this difference may contribute to the discrepancy between their results and ours, we have no definitive explanation.

Third, we have already noted that a pattern of dependence on the environmental stability that corresponds closely to the consensus view is observed in a contiguous environment with symmetric individual learning and mutation among the genotypic values (Table 2c). Namely, the GE strategy, the SE strategy, and the G strategy are favored by natural selection when environmental changes occur at short, intermediate, and long intervals, respectively. Here, it is possible to regard GE as an "evolved" strategy in which the mature phenotype is acquired by individual learning. On the other hand, the initial phenotype is fixed genetically, which constrains the range of mature phenotypes, so that GE is clearly not a "blank slate" strategy.

But Table 2a, 2b, 2c, and 2d shows that this combination of conditions—contiguous environment, symmetric individual learning, and mutable genotypic values—is the only one that produces this pattern (consistent with the consensus view). For example, just changing the environment from contiguous to distant leads to the highly disparate prediction that the GE strategy is fixed for any value of  $l > 1$ , with extinction occurring if  $l = 1$  (Table 2d). When the environment is overlapping, the results are again quite different (Table 2a and 2b). In this case, selection often favors the S strategy (social learning without exploration). In particular, with mutable genotypic values, the S strategy is fixed whatever the interval between environmental changes, and regardless of whether individual learning is symmetric or asymmetric.

As mentioned above, we did not consider values of the interval between environmental changes,  $l$ , greater than 1000 in our numerical work. On the technical side, this choice was mandated by the prohibitively long computing time required to achieve convergence to equilibrium. In some cases, equilibrium had still not been reached after 1,000,000 generations. On the theoretical

side, it is partially justified by the analytical work, which shows that the transitions, if any, between the qualitatively different equilibria mostly occur at relatively small values of  $l$ . For example, we proved in Appendix 3 that, in a nonoverlapping contiguous environment with symmetric learning but no mutation, the threshold between a stable monomorphism of the SE strategy and invasion by the GE strategy is  $l = 4$ .

Finally, we note that incorporation of exogenous costs of learning, such as might be incurred in developing and maintaining a nervous system supportive of learning, may require us to revise our conclusions (Aoki et al. 2005). By way of illustration, let us briefly reexamine the case of the nonoverlapping contiguous environment with symmetric learning and no mutation. In the absence of exogenous costs of learning, our analysis showed that the SE strategy can resist invasion by the GE strategy if  $l \geq 5$ . In particular, if  $l = 5$  and  $w \geq 1$ , the product of the fitnesses over one period of duration  $2l = 10$  is  $\frac{1}{3^6}$  for SE (see (A24c), provided GE is rare) and  $\frac{2^5}{3^{10}}$  for GE (see (A22), provided the genotypic value coincides with an environmental state border). Now, if the exogenous cost of social learning is  $d$  per generation, the above product of the fitnesses for SE must be multiplied by  $(1 - d)^{10}$ . On the other hand, the exogenous cost of individual learning, which is shared by both strategies, can be neglected. Hence, the condition for the GE strategy to invade given  $l = 5$  and  $w \geq 1$  becomes

$$\frac{1}{3^6}(1 - d)^{10} < \frac{2^5}{3^{10}}, \quad (8)$$

or equivalently  $d > 1 - \sqrt[10]{\frac{2^5}{3^4}} \approx 0.089$ .

#### ACKNOWLEDGMENTS

We thank the anonymous reviewers for useful suggestions. Research supported in part by the Morrison Institute for Population and Resource Studies, a grant to the Santa Fe Institute from the James S. McDonnell Foundation 21st Century Collaborative Award Studying Complex Systems, NIH grant GM28016, and Monbukagakusho grant 17102002.

#### LITERATURE CITED

- Akazawa, T., K. Aoki, and O. Bar-Yosef. 1998. Neandertals and modern humans in western Asia. Plenum Press, New York.
- Alvard, M. S. 2003. The adaptive nature of culture. *Evol. Anthropol.* 12:136–149.
- Aoki, K., J. Y. Wakano, and M. W. Feldman. 2005. The emergence of social learning in a temporally changing environment: a theoretical model. *Curr. Anthropol.* 46:334–340.
- Boyd, R., and P. J. Richerson. 1985. *Culture and the evolutionary process*. Univ. Chicago Press, Chicago.
- . 1988. An evolutionary model of social learning: the effect of spatial and temporal variation. Pp. 29–48 in T. Zentall and B. G. Galef Jr, eds. *Social learning*. Erlbaum, Hillsdale, NJ.
- Britton, N. F. 2003. *Essential mathematical biology*. Springer, London.
- Cavalli-Sforza, L. L., and M. W. Feldman. 1981. *Cultural transmission and evolution*. Princeton Univ. Press, Princeton, NJ.
- Doebeli, M., C. Hauert, and T. Killingback. 2004. The evolutionary origin of cooperators and defectors. *Science* 306:859–862.

Durham, W. H. 1991. *Coevolution*. Stanford Univ. Press, Stanford, CA.

Eliassen, S., C. Jørgensen, M. Mangel, and J. Giske. 2007. Exploration or exploitation: life expectancy changes the value of learning in foraging strategies. *Oikos* 116:513–523.

Feldman, M. W., K. Aoki, and J. Kumm. 1996. Individual versus social learning: evolutionary analysis in a fluctuating environment. *Anthropol. Sci.* 104:209–232.

Galef, B. G. Jr. 1988. Imitation in animals: history, definitions, and interpretations of data from the psychological laboratory. Pp. 2–28 in T. Zentall and B. G. Galef Jr, eds. *Social learning*. Erlbaum, Hillsdale, NJ.

Henrich, J., and R. McElreath. 2003. The evolution of cultural evolution. *Evol. Anthropol.* 12:123–135.

Heyes, C. M. 1993. Imitation, culture, and cognition. *Anim. Behav.* 46:999–1010.

Klein, R. 1999. *Human career*, 2nd ed. Univ. of Chicago Press, Chicago, IL.

Laland, K. N., and W. Hoppitt. 2003. Do animals have culture? *Evol. Anthropol.* 12:150–159.

Laland, K. N., J. Odling-Smee, and M. W. Feldman. 2000. Niche construction, biological evolution, and cultural change. *Behav. Brain Sci.* 23:131–175.

Richerson, P. J., and R. Boyd. 2000. Built for speed: Pleistocene climate variation and the origin of human culture. *Pers. Ethol.* 13:1–45.

Rodriguez-Gironés, M. A., and R. A. Vásquez. 1997. Density-dependent patch exploitation and acquisition of environmental information. *Theor. Pop. Biol.* 52:32–42.

Rogers, A. R. 1988. Does biology constrain culture? *Am. Anthropol.* 90:819–831.

Tomasello, M. 1999. The human adaptation for culture. *Ann. Rev. Anthropol.* 28:509–529.

Wakano, J. Y., and K. Aoki. 2006. A mixed strategy model for the emergence and intensification of social learning in a periodically changing environment. *Theor. Pop. Biol.* 70:486–497.

Wakano, J. Y., K. Aoki, and M. W. Feldman. 2004. Evolution of social learning: a mathematical analysis. *Theor. Pop. Biol.* 66:249–258.

Whiten, A., and R. Ham. 1992. On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. Pp. 239–283 in J. S. Rosenblatt, R. A. Hinde, E. Shaw, and C. Beer, eds. *Advances in the study of behavior*, vol. 21. Academic Press, New York.

Associate Editor: J. Wolf

## Appendix 1

### THE THREE-LOCUS MODEL RECURSIONS

Here we give the recursions in their most general form. Some simplifying assumptions are made in the text and following appendices.

Let  $\eta_{t-1}(z^*)$  be the distribution of surviving phenotypes  $z^*$  in generation  $t - 1$ . As defined by equation (1), the distribution of initial phenotypes  $x$  among  $g, k$  organisms in generation  $t$  is

$$\psi_t(x | g, k) = (1 - k)\delta_{xg} + k\eta_{t-1}(x). \quad (A1)$$

Next let  $\lambda(z | x, b)$  be the individual learning function for  $b$  organisms whose initial phenotype is  $x$ . Then, the distribution of mature phenotypes  $z$  among  $g, k, b$  organisms in generation  $t$  is

$$\xi_t(z | g, k, b) = \sum_x \lambda(z | x, b)\psi_t(x | g, k). \quad (A2)$$

The fitness of a  $g, k, b$  organism in environmental state  $y$  in generation  $t$  is

$$w_t(g, k, b | y) = \sum_{z=y-w}^{y+w} \xi_t(z | g, k, b). \quad (A3)$$

Hence, the recursion in the frequency  $\phi_t(g, k, b)$  of haplotype  $g, k, b$  in environmental state  $y$  is

$$\phi_{t+1}(g, k, b) = w_t(g, k, b | y)\phi_t(g, k, b)/\bar{w}_t, \quad (A4)$$

where

$$\bar{w}_t = \sum_g \sum_k \sum_b w_t(g, k, b | y)\phi_t(g, k, b) \quad (A5)$$

is the mean fitness in environmental state  $y$  in generation  $t$ . Finally, the distribution of surviving phenotypes  $z^*$  in environmental state  $y$  in generation  $t$  can be written as

$$\eta_t(z^*) = \sum_g \sum_k \sum_b \xi_t(z^* | g, k, b)\phi_t(g, k, b)/\bar{w}_t. \quad (A6)$$

if  $|z^* - y| \leq w$ , and 0 otherwise.

## Appendix 2

### EQUILIBRIUM IN AN OVERLAPPING ENVIRONMENT WITH SYMMETRIC LEARNING

Let there be two alleles at each locus with  $g = g_1$  or  $g_2$  (where  $g_1 < g_2$ ),  $k = 0$  or 1, and  $b = 0$  or 1. Individual learning is symmetric (eq. 2). The competing generic strategies are G (genetic, innate;  $k = 0, b = 0$ ), GE (genetic explorer;  $k = 0, b = 1$ ), S (social learner;  $k = 1, b = 0$ ), and SE (social learner explorer;  $k = 1, b = 1$ ). Assume overlapping environments so that some mature phenotypes are viable in both states  $y_1$  and  $y_2$  (where  $y_1 < y_2$ ), i.e.,  $y_2 - w \leq y_1 + w$ .

Using Kronecker's  $\delta$ , the individual learning functions for the two alleles  $b = 0$  and 1 can be expressed as

$$\lambda(z | x, 0) = \delta_{zx}, \lambda(z | x, 1) = \frac{1}{3}(\delta_{z,x-1} + \delta_{zx} + \delta_{z,x+1}), \quad (A7)$$

respectively.

To prove the existence of an equilibrium (denoted by the caret above the symbol) in which the haplotype frequencies and phenotype distribution are constant in time, we set

$$\hat{\eta}(z^*) = \sum_{\sigma=y_2-w}^{y_1+w} \beta_\sigma \delta_{z^*\sigma}, \text{ where } \sum_{\sigma=y_2-w}^{y_1+w} \beta_\sigma = 1 \quad (A8)$$

and demonstrate consistency. Substituting (A8) in (A1) gives

$$\hat{\psi}(x | g, 0) = \delta_{xg}, \hat{\psi}(x | g, 1) = \sum_{\sigma=y_2-w}^{y_1+w} \beta_\sigma \delta_{x\sigma}. \quad (A9)$$

Hence, from (A2)

$$\hat{\xi}(z | g, 0, 0) = \sum_x \lambda(z | x, 0) \hat{\psi}(x | g, 0) = \sum_x \delta_{zx} \delta_{xg} = \delta_{zg}, \quad (\text{A10a})$$

$$\begin{aligned} \hat{\xi}(z | g, 0, 1) &= \sum_x \lambda(z | x, 1) \hat{\psi}(x | g, 0) \\ &= \frac{1}{3} \sum_x (\delta_{z,x-1} + \delta_{zx} + \delta_{z,x+1}) \delta_{xg} \\ &= \frac{1}{3} (\delta_{z,g-1} + \delta_{zg} + \delta_{z,g+1}), \end{aligned} \quad (\text{A10b})$$

$$\begin{aligned} \hat{\xi}(z | g, 1, 0) &= \sum_x \lambda(z | x, 0) \hat{\psi}(x | g, 1) = \sum_x \delta_{zx} \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} \delta_{x\sigma} \\ &= \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} \sum_x \delta_{zx} \delta_{x\sigma} \\ &= \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} \delta_{z\sigma} = \hat{\eta}(z), \end{aligned} \quad (\text{A10c})$$

$$\begin{aligned} \hat{\xi}(z | g, 1, 1) &= \sum_x \lambda(z | x, 1) \hat{\psi}(x | g, 1) \\ &= \frac{1}{3} \sum_x (\delta_{z,x-1} + \delta_{zx} + \delta_{z,x+1}) \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} \delta_{x\sigma} \\ &= \frac{1}{3} \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} (\delta_{z,\sigma-1} + \delta_{z\sigma} + \delta_{z,\sigma+1}). \end{aligned} \quad (\text{A10d})$$

Thus from (A3), (A10a), and (A10b) the fitnesses of the G and GE strategists are

$$\begin{aligned} \hat{w}(g, 0, 0 | y) &= \sum_{z=y-w}^{y+w} \hat{\xi}(z | g, 0, 0) = \sum_{z=y-w}^{y+w} \delta_{zg} \\ &= 1 \text{ if } |g - y| \leq w, 0 \text{ otherwise} \end{aligned} \quad (\text{A11a})$$

$$\begin{aligned} \hat{w}(g, 0, 1 | y) &= \sum_{z=y-w}^{y+w} \hat{\xi}(z | g, 0, 1) \\ &= \frac{1}{3} \left( \sum_{z=y-w}^{y+w} \delta_{z,g-1} + \sum_{z=y-w}^{y+w} \delta_{zg} + \sum_{z=y-w}^{y+w} \delta_{z,g+1} \right) \\ &= 1 \text{ if } |g - y| \leq w - 1, \frac{2}{3} \text{ if } |g - y| = w, \\ &= \frac{1}{3} \text{ if } |g - y| = w + 1, 0 \text{ otherwise.} \end{aligned} \quad (\text{A11b})$$

For the S strategists (A3) and (A10c) give

$$\begin{aligned} \hat{w}(g, 1, 0 | y) &= \sum_{z=y-w}^{y+w} \hat{\xi}(z | g, 1, 0) = \sum_{z=y-w}^{y+w} \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} \delta_{z\sigma} \\ &= \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} \sum_{z=y-w}^{y+w} \delta_{z\sigma}. \end{aligned}$$

Because  $y_1 - w < y_2 - w < y_1 + w < y_2 + w$ , we have  $\sum_{z=y-w}^{y+w} \delta_{z\sigma} = 1$  for each  $\sigma$ . Hence

$$\hat{w}(g, 1, 0 | y) = \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} = 1. \quad (\text{A11c})$$

Similarly, for the SE strategists (A3) and (A10d) give

$$\begin{aligned} \hat{w}(g, 1, 1 | y) &= \sum_{z=y-w}^{y+w} \hat{\xi}(z | g, 1, 1) \\ &= \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} \frac{1}{3} \left( \sum_{z=y-w}^{y+w} \delta_{z,\sigma-1} + \sum_{z=y-w}^{y+w} \delta_{z\sigma} + \sum_{z=y-w}^{y+w} \delta_{z,\sigma+1} \right). \end{aligned}$$

Noting that  $\sum_{z=y_2-w}^{y_2+w} \delta_{z,\sigma-1} = 0$  when  $\sigma = y_2 - w$  and  $\sum_{z=y_1-w}^{y_1+w} \delta_{z,\sigma+1} = 0$  when  $\sigma = y_1 + w$ , regardless of the value of  $g$ , we infer

$$\hat{w}(g, 1, 1 | y) < \sum_{\sigma=y_2-w}^{y_1+w} \beta_{\sigma} = 1. \quad (\text{A11d})$$

(Note that the functions  $\psi$ ,  $\xi$ , and  $w$  do not depend on  $t$ .)

Haplotypes with fitness 1 in both states will persist, whereas haplotypes of lower fitness will be selected out. Hence, (1) the G strategy haplotype  $g, 0, 0$  persists if  $|g - y_1| \leq w$  and  $|g - y_2| \leq w$ , i.e., if  $y_2 - w \leq g \leq y_1 + w$ , but is otherwise selected out; (2) the GE strategy haplotype  $g, 0, 1$  persists if  $|g - y_1| \leq w - 1$  and  $|g - y_2| \leq w - 1$ , i.e., if  $y_2 - w + 1 \leq g \leq y_1 + w - 1$ , but is otherwise selected out; (3) the S strategy haplotypes  $g_1, 1, 0$  and  $g_2, 1, 0$  persist; (4) the SE strategy haplotypes  $g_1, 1, 1$  and  $g_2, 1, 1$  are selected out.

Thus, in the least restrictive case (i.e.,  $g < y_2 - w$  or  $y_1 + w < g$  for  $g = g_1, g_2$ ), only the two S haplotypes  $g_1, 1, 0$  and  $g_2, 1, 0$  exist at equilibrium. When the conditions on  $g$  are made more stringent, some or all of the G or GE haplotypes may coexist. Note that the frequencies of persisting haplotypes will be constant at equilibrium, but the values depend on the initial conditions.

Lastly, we show the existence of an equilibrium phenotypic distribution that is constant,  $\hat{\eta}(z^*)$ . In other words, we show that  $\eta_{t-1}(z^*) = \hat{\eta}(z^*)$  implies  $\eta_t(z^*) = \hat{\eta}(z^*)$ . The mean fitness  $\hat{w} = 1$  because each persisting haplotype has a fitness of 1, so we need to only evaluate the numerator of (A6). There are three cases. First suppose  $\phi_i(g_i, 1, 0) = \hat{\phi}(g_i, 1, 0)$  for  $i = 1, 2$ , where  $\hat{\phi}(g_1, 1, 0) + \hat{\phi}(g_2, 1, 0) = 1$ . Then, substituting  $\xi_i(z | g_i, 1, 0) = \hat{\eta}(z)$  for  $i = 1, 2$  (eq. A10c) in (A6), the distribution of surviving phenotypes in generation  $t$  will be

$$\eta_t(z^*) = \hat{\eta}(z^*) \hat{\phi}(g_1, 1, 0) + \hat{\eta}(z^*) \hat{\phi}(g_2, 1, 0) = \hat{\eta}(z^*), \quad (\text{A12a})$$

and we are done. Second, if  $\hat{\phi}(g_1, 1, 0) + \hat{\phi}(g_2, 1, 0) + \hat{\phi}(g_1, 0, 0) + \hat{\phi}(g_2, 0, 0) = 1$ , then

$$\hat{\eta}(z^*) = \hat{\eta}(z^*)\hat{\phi}(g_1, 1, 0) + \hat{\eta}(z^*)\hat{\phi}(g_2, 1, 0) + \delta_{z^*g_1}\hat{\phi}(g_1, 0, 0) + \delta_{z^*g_2}\hat{\phi}(g_2, 0, 0),$$

whence

$$\hat{\eta}(z^*) = \frac{\delta_{z^*g_1}\hat{\phi}(g_1, 0, 0) + \delta_{z^*g_2}\hat{\phi}(g_2, 0, 0)}{\hat{\phi}(g_1, 0, 0) + \hat{\phi}(g_2, 0, 0)}. \tag{A12b}$$

Third, if  $\hat{\phi}(g_1, 1, 0) + \hat{\phi}(g_2, 1, 0) + \hat{\phi}(g_1, 0, 0) + \hat{\phi}(g_2, 0, 0) + \hat{\phi}(g_1, 0, 1) + \hat{\phi}(g_2, 0, 1) = 1$ , then

$$\begin{aligned} &\hat{\eta}(z^*)(1 - \hat{\phi}(g_1, 1, 0) - \hat{\phi}(g_2, 1, 0)) \\ &= \delta_{z^*g_1}\hat{\phi}(g_1, 0, 0) + \delta_{z^*g_2}\hat{\phi}(g_2, 0, 0) \\ &\quad + \frac{1}{3}(\delta_{z^*,g_1-1} + \delta_{z^*g_1} + \delta_{z^*,g_1+1})\hat{\phi}(g_1, 0, 1) \\ &\quad + \frac{1}{3}(\delta_{z^*,g_2-1} + \delta_{z^*g_2} + \delta_{z^*,g_2+1})\hat{\phi}(g_2, 0, 1). \end{aligned} \tag{A12c}$$

### Appendix 3

#### EQUILIBRIUM IN A CONTIGUOUS ENVIRONMENT WITH SYMMETRIC LEARNING

We retain all assumptions made in Appendix 2, except that the environment is now nonoverlapping contiguous, i.e.,  $y_1 + w + 1 = y_2 - w$ . Consider an equilibrium at which only the SE strategists are present. Then clearly the two haplotypes  $g_1, 1, 1$  and  $g_2, 1, 1$  have the same fitness, whence their frequencies are constant and can be written as  $\hat{\phi}(g_1, 1, 1)$  and  $\hat{\phi}(g_2, 1, 1)$ , respectively.

Substitution of (2) and (A1) in (A2) yields

$$\hat{\xi}_t(z | g, k, b) = (1 - k)\lambda(z | g, b) + k \frac{1}{2b + 1} \sum_{x=z-b}^{z+b} \hat{\eta}_{t-1}(x), \tag{A13}$$

which simplifies to

$$\hat{\xi}_t(z | g, 1, 1) = \frac{1}{3} \sum_{x=z-1}^{z+1} \hat{\eta}_{t-1}(x), \tag{A14}$$

for haplotype  $g, 1, 1$ .

In particular, in the first generation after a change of state from  $y_1$  to  $y_2$ , we have from (A14)

$$\hat{\xi}_1(z | g, 1, 1) = \frac{1}{3} \sum_{x=z-1}^{z+1} \hat{\eta}_0(x), \tag{A15}$$

where the subscript 0 denotes the generation before the change of state, and  $\hat{\eta}_0(x) = 0$  for  $x \geq y_2 - w$ . Hence,

$$\begin{aligned} \hat{\xi}_1(y_2 - w | g, 1, 1) &= \frac{1}{3}\hat{\eta}_0(y_2 - w - 1) = \frac{1}{3}\hat{\eta}_0(y_1 + w), \\ \hat{\xi}_1(z | g, 1, 1) &= 0 \quad \text{for } z > y_2 - w. \end{aligned} \tag{A16}$$

Thus the fitnesses of these SE strategists, whether  $g = g_1$  or  $g_2$ , are

$$\hat{w}_1(g, 1, 1 | y_2) = \frac{1}{3} \sum_{z=y_2-w}^{y_2+w} \hat{\xi}_1(z | g, 1, 1) = \frac{1}{3}\hat{\eta}_0(y_1 + w). \tag{A17}$$

Substituting (A16) and (A17) in (A6) yields

$$\hat{\eta}_1(y_2 - w) = 1. \tag{A18}$$

While the environment remains in state  $y_2$  ( $2 \leq t \leq l$ ), equation (A14) applies with  $\hat{\eta}_{t-1}(x) = 0$  for  $|x - y_2| > w$ . Hence, the fitnesses are

$$\hat{w}_t(g, 1, 1 | y_2) = \frac{1}{3} \sum_{z=y_2-w}^{y_2+w} \sum_{x=z-1}^{z+1} \hat{\eta}_{t-1}(x) \tag{A19a}$$

$$= \frac{1}{3}\hat{\eta}_{t-1}(y_2) \text{ if } w = 0, \tag{A19b}$$

$$= 1 - \frac{1}{3}\hat{\eta}_{t-1}(y_2 - w) - \frac{1}{3}\hat{\eta}_{t-1}(y_2 + w) \text{ if } w \geq 1. \tag{A19c}$$

Moreover, when the haplotype frequencies are at equilibrium, (A6), (A14), and (A19a) entail that the distribution of surviving phenotypes satisfies the recursion

$$\hat{\eta}_t(z^*) = \frac{\sum_{x=z^*-1}^{z^*+1} \hat{\eta}_{t-1}(x)}{\sum_{u=y_2-w}^{y_2+w} \sum_{x=u-1}^{u+1} \hat{\eta}_{t-1}(x)} \tag{A20}$$

for  $|z^* - y_2| \leq w$ . In particular,  $\hat{\eta}_t(y_2) = 1$  and  $\hat{w}_t(g, 1, 1 | y_2) = \frac{1}{3}$  for  $1 \leq t \leq l$  if  $w = 0$ . The relative abundances of the surviving phenotypes can be computed using (A18) and (A20) and are shown in Table 1a and 1b.

Clearly, a mirror-image pattern is repeated when environment reverts to state  $y_1$ .

Let us now introduce the GE strategy haplotypes  $g_1, 0, 1$  and  $g_2, 0, 1$  at low frequencies. Unless  $g_1 = y_1 + w$  or  $g_2 = y_2 - w$ , both haplotypes will be selected out. We set  $g_1 = y_1 + w$ , and compute the fitness of haplotype  $g_1, 0, 1$ . Assuming  $w \geq 1$ , the fitnesses of haplotype  $g_1, 0, 1$  in environmental states  $y_1$  and  $y_2$  are

$$w_t(g_1, 0, 1 | y_1) = \frac{2}{3}, w_t(g_1, 0, 1 | y_2) = \frac{1}{3}. \tag{A21}$$

Hence, the product of the fitnesses over one period of duration  $2l$  is

$$\hat{w}_{GE}(g_1, 0, 1) = \left(\frac{1}{3}\right)^l \left(\frac{2}{3}\right)^l = \frac{2^l}{3^{2l}}. \tag{A22}$$

On the other hand, the fitnesses of the SE strategists ( $g = g_1$  or

$g_2$ ) when GE strategists are rare can be approximated by

$$\begin{aligned} w_1(g, 1, 1 | y_2) &= \frac{1}{3} \hat{\eta}_0(y_1 + w) \equiv \frac{1}{3} \hat{\eta}_t(y_1 + w) \\ w_t(g, 1, 1 | y_1) &= \frac{1}{3} \hat{\eta}_0(y_2 - w) \equiv \frac{1}{3} \hat{\eta}_t(y_2 - w), \end{aligned} \quad (\text{A23a})$$

and from (A19c)

$$\begin{aligned} w_t(g, 1, 1 | y_2) &= 1 - \frac{1}{3} \hat{\eta}_{t-1}(y_2 - w) - \frac{1}{3} \hat{\eta}_{t-1}(y_2 + w) \\ w_t(g, 1, 1 | y_1) &= 1 - \frac{1}{3} \hat{\eta}_{t-1}(y_1 + w) - \frac{1}{3} \hat{\eta}_{t-1}(y_1 - w) \end{aligned} \quad (\text{A23b})$$

for  $2 \leq t \leq l$ .

Referring to Table 1a or 1b, and using (A23a) and (A23b), the product of the fitnesses when  $l = 3$  is

$$\tilde{w}_{SE}(g, 1, 1) = \left[ \frac{1}{3} \cdot \frac{2}{5} \left( 1 - \frac{1}{3} \cdot 1 \right) \left( 1 - \frac{1}{3} \cdot \frac{1}{2} \right) \right]^2 = \frac{2^2}{3^6}. \quad (\text{A24a})$$

Comparing (A24a) with (A22),  $\tilde{w}_{SE}(g_1, 1, 1) = \tilde{w}_{SE}(g_2, 1, 1) < \tilde{w}_{GE}(g_1, 0, 1)$  when  $l = 3$ , and the equilibrium is unstable to invasion by the GE strategy haplotype  $g_1, 0, 1$ .

For  $l \geq 4$  the entries of Table 1a ( $w = 1$ ) and Table 1b ( $w \geq 2$ ) differ. Here we compute the fitnesses of SE organisms assuming  $w = 1$ , but we get identical results when  $w \geq 2$ . Then

$$\begin{aligned} \tilde{w}_{SE}(g, 1, 1) &= \left[ \frac{1}{3} \cdot \frac{4}{12} \left( 1 - \frac{1}{3} \cdot 1 \right) \left( 1 - \frac{1}{3} \cdot \frac{1}{2} \right) \right. \\ &\quad \left. \times \left( 1 - \frac{1}{3} \cdot \frac{2}{5} - \frac{1}{3} \cdot \frac{1}{5} \right) \right]^2 = \frac{2^4}{3^8}, \end{aligned} \quad (\text{A24b})$$

whence from (A22)  $\tilde{w}_{SE}(g_1, 1, 1) = \tilde{w}_{SE}(g_2, 1, 1) = \tilde{w}_{GE}(g_1, 0, 1)$ , and the eigenvalue is one.

For  $l = 5$

$$\begin{aligned} \tilde{w}_{SE}(g, 1, 1) &= \left[ \frac{1}{3} \cdot \frac{9}{29} \left( 1 - \frac{1}{3} \cdot 1 \right) \right. \\ &\quad \times \left( 1 - \frac{1}{3} \cdot \frac{1}{2} \right) \left( 1 - \frac{1}{3} \cdot \frac{2}{5} - \frac{1}{3} \cdot \frac{1}{5} \right) \\ &\quad \left. \times \left( 1 - \frac{1}{3} \cdot \frac{4}{12} - \frac{1}{3} \cdot \frac{3}{12} \right) \right]^2 = \frac{1}{3^6} \end{aligned} \quad (\text{A24c})$$

so that  $\tilde{w}_{SE}(g_1, 1, 1) = \tilde{w}_{SE}(g_2, 1, 1) > \tilde{w}_{GE}(g_1, 0, 1)$ , and the equilibrium with only SE present is stable.

## Appendix 4

### EQUILIBRIUM IN A CONTIGUOUS ENVIRONMENT WITH ASYMMETRIC LEARNING

Assume a nonoverlapping contiguous environment and asymmetric individual learning (eq. 3). For the special case of  $w = 1$ , we investigate the properties of the equilibrium at which the SE strategy is fixed and also its stability to invasion by the S strategy. Under these conditions, the viable phenotypes in environmental

state  $y_1$  are  $y_1 - 1$ ,  $y_1$ , and  $y_1 + 1$ , whereas the viable phenotypes in environmental state  $y_2$  are  $y_2 - 1$ ,  $y_2$ , and  $y_2 + 1$ . Moreover, phenotypes  $y_1 + 1$  and  $y_2 - 1$  differ by one unit.

Immediately after a change of state from  $y_1$  to  $y_2$  ( $y_1 < y_2$ ), we have using (3a), (A1), and (A2)

$$\begin{aligned} \hat{\xi}_1(y_2 - 1 | g, 1, 1) &= \frac{1}{3} \hat{\eta}_0(y_1) + \frac{1}{3} \hat{\eta}_0(y_1 + 1), \\ \hat{\xi}_1(y_2 | g, 1, 1) &= \frac{1}{3} \hat{\eta}_0(y_1 + 1), \\ \hat{\xi}_1(y_2 + 1 | g, 1, 1) &= 0. \end{aligned} \quad (\text{A25})$$

Hence, from (A3) the fitness of haplotype  $g, 1, 1$  is

$$\hat{w}_1(g, 1, 1 | y_2) = \frac{1}{3} \hat{\eta}_0(y_1) + \frac{2}{3} \hat{\eta}_0(y_1 + 1), \quad (\text{A26})$$

Substitution of (A25) and (A26) (which hold for both  $g = g_1$  and  $g_2$ ) in (A6) yields

$$\begin{aligned} \hat{\eta}_1(y_2 - 1) &= \frac{\hat{\eta}_0(y_1) + \hat{\eta}_0(y_1 + 1)}{\hat{\eta}_0(y_1) + 2\hat{\eta}_0(y_1 + 1)}, \\ \hat{\eta}_1(y_2) &= \frac{\hat{\eta}_0(y_1 + 1)}{\hat{\eta}_0(y_1) + 2\hat{\eta}_0(y_1 + 1)}, \\ \hat{\eta}_1(y_2 + 1) &= 0. \end{aligned} \quad (\text{A27})$$

Similarly, while the environment remains in state  $y_2$  ( $2 \leq t \leq l$ ),

$$\begin{aligned} \hat{\xi}_t(y_2 - 1 | g, 1, 1) &= \frac{1}{3} \hat{\eta}_{t-1}(y_2 - 1) + \frac{1}{3} \hat{\eta}_{t-1}(y_2) \\ &\quad + \frac{1}{3} \hat{\eta}_{t-1}(y_2 + 1) = \frac{1}{3}, \\ \hat{\xi}_t(y_2 | g, 1, 1) &= \frac{1}{3} \hat{\eta}_{t-1}(y_2 - 1) + \frac{1}{3} \hat{\eta}_{t-1}(y_2) \\ &\quad + \frac{1}{3} \hat{\eta}_{t-1}(y_2 + 1) = \frac{1}{3}, \\ \hat{\xi}_t(y_2 + 1 | g, 1, 1) &= \frac{1}{3} \hat{\eta}_{t-1}(y_2 - 1) + \frac{1}{3} \hat{\eta}_{t-1}(y_2) \\ &\quad + \frac{1}{3} \hat{\eta}_{t-1}(y_2 + 1) = \frac{1}{3}. \end{aligned} \quad (\text{A28})$$

Hence,

$$\hat{w}_t(g, 1, 1 | y_2) = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1, \quad (\text{A29})$$

and

$$\hat{\eta}_t(y_2 - 1) = \hat{\eta}_t(y_2) = \hat{\eta}_t(y_2 + 1) = \frac{1}{3}. \quad (\text{A30})$$

Next, the fitness of the S strategy haplotype  $g, 1, 0$  introduced at low frequency is approximately

$$w_1(g, 1, 0 | y_2) = \hat{\eta}_0(y_1 + 1) \quad (\text{A31})$$

immediately after a change of state from  $y_1$  to  $y_2$ , and

$$w_t(g, 1, 0 | y_2) = \hat{\eta}_{t-1}(y_2 - 1) + \hat{\eta}_{t-1}(y_2) + \hat{\eta}_{t-1}(y_2 + 1) = 1. \quad (\text{A32})$$

in generations  $2 \leq t \leq l$ .

We need to distinguish the two cases  $l = 1$  and  $l \geq 2$ . In the former case, the environment changes every generation, which requires repeated application of (A27). Moreover, symmetry of the two environmental states dictates that at equilibrium  $\hat{\eta}_1(y_2 - 1) = \hat{\eta}_0(y_1 + 1)$  and  $\hat{\eta}_1(y_2) = \hat{\eta}_0(y_1)$ . Hence, the solution of (A27) is

$$\hat{\eta}_0(y_1) = \frac{3 - \sqrt{5}}{2}, \hat{\eta}_0(y_1 + 1) = \frac{\sqrt{5} - 1}{2}. \tag{A33}$$

Substituting (A33) into (A26) and (A31) gives

$$\begin{aligned} \hat{w}_1(g, 1, 1 | y_2) &= \frac{\sqrt{5} + 1}{6}, \\ w_1(g, 1, 0 | y_2) &= \frac{\sqrt{5} - 1}{2}, \end{aligned} \tag{A34}$$

respectively. Invoking symmetry again, the product of the fitnesses over one environmental cycle (of length 2) is

$$\tilde{w}_{SE} = \left( \frac{\sqrt{5} + 1}{6} \right)^2 \tag{A35}$$

for the SE strategists, and

$$\tilde{w}_S = \left( \frac{\sqrt{5} - 1}{2} \right)^2 \tag{A36}$$

for the S strategists. Because  $\tilde{w}_{SE} < \tilde{w}_S$ , the S strategy will invade.

Next, in the case of  $l \geq 2$ , (A29) and (A30) indicate that the fitness of haplotype  $g, 1, 1$  and the phenotype distribution are both constant after the second generation in state  $y_2$  (i.e., for  $2 \leq t \leq l$ ). Because generations 0 and  $l$  are equivalent, we obtain at equilibrium

$$\hat{\eta}_0(y_2 - 1) = \hat{\eta}_0(y_2) = \hat{\eta}_0(y_2 + 1) = \frac{1}{3}, \tag{A37}$$

and

$$\hat{w}_1(g, 1, 1 | y_2) = \frac{1}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{3} = \frac{1}{3}. \tag{A38}$$

Thus the product of the fitnesses over one cycle of duration  $2l$  is

$$\tilde{w}_{SE} = \left( \frac{1}{3} \right)^2 \cdot 1^{2(l-1)} = \frac{1}{9}. \tag{A39}$$

On the other hand, the fitness of the rare S strategists is approximately

$$w_1(g, 1, 0 | y_2) = \hat{\eta}_0(y_1 + 1) = \frac{1}{3}, \tag{A40a}$$

and

$$w_t(g, 1, 0 | y_2) = \hat{\eta}_{t-1}(y_2 - 1) + \hat{\eta}_{t-1}(y_2) + \hat{\eta}_{t-1}(y_2 + 1) = 1 \tag{A40b}$$

for  $2 \leq t \leq l$ . Hence,

$$\tilde{w}_S = \left( \frac{1}{3} \right)^2 \cdot 1^{2(l-1)} = \frac{1}{9}. \tag{A41}$$

Comparison of (A39) and (A41) shows that  $\tilde{w}_{SE} = \tilde{w}_S$ .

Finally, we sketch a proof for the special case of  $w = 2$ . Immediately after a change of state from  $y_1$  to  $y_2$  ( $y_1 < y_2$ ), the phenotype distribution at equilibrium when the SE strategy is fixed satisfies

$$\begin{aligned} \hat{\eta}_1(y_2 - 2) &= \frac{\hat{\eta}_0(y_1 + 1) + \hat{\eta}_0(y_1 + 2)}{\hat{\eta}_0(y_1 + 1) + 2\hat{\eta}_0(y_1 + 2)}, \\ \hat{\eta}_1(y_2 - 1) &= \frac{\hat{\eta}_0(y_1 + 2)}{\hat{\eta}_0(y_1 + 1) + 2\hat{\eta}_0(y_1 + 2)}, \\ \hat{\eta}_1(y_2) &= 0, \\ \hat{\eta}_1(y_2 + 1) &= 0, \\ \hat{\eta}_1(y_2 + 2) &= 0. \end{aligned} \tag{A42}$$

Subsequently, while the environment remains in state  $y_2$  ( $2 \leq t \leq l$ ), we have

$$\begin{aligned} \begin{pmatrix} \hat{\eta}_t(y_2 - 2) \\ \hat{\eta}_t(y_2 - 1) \\ \hat{\eta}_t(y_2) \\ \hat{\eta}_t(y_2 + 1) \end{pmatrix} &= \begin{pmatrix} 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{pmatrix} \\ &\times \begin{pmatrix} \hat{\eta}_{t-1}(y_2 - 2) \\ \hat{\eta}_{t-1}(y_2 - 1) \\ \hat{\eta}_{t-1}(y_2) \\ \hat{\eta}_{t-1}(y_2 + 1) \end{pmatrix}. \end{aligned} \tag{A43}$$

Clearly, after many generations of stasis in state  $y_2$ , the equilibrium phenotype distribution will approach  $\hat{\eta}_\infty(y_2 - 2) = 0$  and  $\hat{\eta}_\infty(y_2 - 1) = \hat{\eta}_\infty(y_2) = \hat{\eta}_\infty(y_2 + 1) = 1/3$ . Hence, if  $l$  is sufficiently large, then in the generation immediately after a change of state back to  $y_1$ , the fitness of the SE strategists is approximately  $1/9$ , whereas the fitness of the S strategists is negligibly small.